# Part II

- Finite Element Methods for Elliptic Equations

  1. 1D model problem

     - Variational (Weak) formulation
     - Introduction to Sobolev spaces ($L^2, H^1$ and $H_0^1$, etc.)
     - Galerkin methods
     - Finite Element Methods using piecewise linear functions
     - Error estimation

  2. 2D Poisson equation

     - Sobolev spaces ($L^2(\Omega), H^1(\Omega)$ and $H_0^1(\Omega)$, etc.)
     - Finite Element Methods using piecewise bi-linear functions
     - The best error approximation: Geometric interpretation
     - The Neumann problem: Natural and essential boundary conditions

- FEM Programming

- Finite Difference/Finite Element Methods for Parabolic Equations
- Discontinuous Galerkin Methods for Hyperbolic Equations (stage/2010/DG.pdf)

# FEM for Elliptic Equations

## 1D model problem

Consider the following boundary value problem (BVP or SP, Strong Problem):

$$\begin{cases} -u''(x) = f(x), \quad x \in (0,1), \\ u(0) = u(1) = 0, \end{cases}$$

where $f$ is a given continuous function.

- By integrating twice, we can see that this problem has a unique solution.

# Variational formulation of the model problem

The Minimization Problem (MP): Find $u \in V$, such that $\mathcal{F}(u) \leqslant \mathcal{F}(v), \ \forall v \in V$, where $\mathcal{F}$ is the linear functional $V \to R$, defined as:

$$\mathcal{F}(v) = \frac{1}{2}(v', v') - (f, v), \ \forall v \in V,$$

- $V = \{v : v$ and $v'$ are square integrable on $[0, 1]$, and $v(0) = v(1) = 0\}$;
- $(v, w) = \int_0^1 v(x)w(x)dx$, for all real valued piecewise continuous functions $v, w$;

• (MP) corresponds to the "Principle of Minimum Potential Energy" in Mechanics.

The Variational (Weak) Problem (VP): Find $u \in V$, such that

$$(u', v') = (f, v), \forall v \in V.$$

- (VP) corresponds to the "Principle of Virtual Work" in Mechanics.

## Relationship between (SP), (MP), and (VP)

The solution of (SP) is also a solution of (VP)

**Proof**:

- Multiplying the equation $-u'' = f$ by an arbitrary function $v \in V$;
- Integrating over $(0, 1)$ which gives:

$$(-u'', v) = (f, v), \ \forall v \in V;$$

- Applying the integration by parts in the left-hand side and using the fact that $v(0) = v(1) = 0$ to get:

$$-(u'', v) = (u', v') - u'(1)v(1) + u'(0)v(0) = (u', v');$$

- Finally, we conclude that: $(u', v') = (f, v)$ for all $v \in V$ , i.e. $u$ is a solution of (VP).

The solution of (VP) is also a solution of (MP)

**Proof**:

- Let $u$ be a solution to (VP), let $v \in V$ and set $w = v - u$ so that $v = u + w, w \in V$.
- Then

$$\begin{aligned}
\mathcal{F}(v) &= \mathcal{F}(u+w) = \frac{1}{2}(u'+w', u'+w') - (f, u+w) \\
&= \frac{1}{2}(u', u') - (f, u) + (u', w') - (f, w) + \frac{1}{2}(w', w') \\
&= \mathcal{F}(u) + \frac{1}{2}(w', w') \\
&\geqslant \mathcal{F}(u);
\end{aligned}$$

- Thus, $u$ is the minimizer of $\mathcal{F}(v)$ (i.e., $u$ is a solution of (MP)).

<u>The solution of (MP) is also a solution of (VP)</u>

**Proof**:

- Let $u$ be a solution to (MP), then for any real number $\alpha$ and any $v \in V$, we have that $u + \alpha v \in V$, which implies:

$$\mathcal{F}(u) \leqslant \mathcal{F}(u + \alpha v);$$

- Thus the differentiable function $g(\alpha) = \mathcal{F}(u + \alpha v)$ has a minimum at $\alpha = 0$ and hence $g'(0) = 0$.
- By direct calculation:

$$g(\alpha) = \frac{1}{2}(u', u') + \alpha(u', v') + \frac{\alpha^2}{2}(v', v') - (f, u) - \alpha(f, v),$$

and thus

$$g'(\alpha) = (u', v') + \alpha(v', v') - (f, v);$$

- Using $g'(0) = 0$ and $g'(0) = (u', v') - (f, v)$, results in:

$$(u', v') = (f, v), \ \forall v \in V$$

i.e. $u$ is a solution of (VP).

<u>The solution of (VP) is also a solution of (SP)</u>

**Proof**:

- Let $u \in V$ be a solution of (VP), then:

$$\int_0^1 u'(x)v'(x)dx - \int_0^1 f(x)v(x)dx = 0$$

- Assuming that $u''$ exists and is continuous (regularity assumption);
- Integrating the first term and using $v(0) = v(1) = 0$, we have

$$\int_0^1 (u''(x) + f(x))v(x)dx = 0, \ \forall v \in V.$$

- By the assumption that $u'' + f$ is continuous, the above relation can only hold if:

$$-u''(x) = f(x), \ \forall x \in [0, 1].$$

So $u$ is a solution of (SP).

• The above results mean the equivalence between (SP), (VP), and (MP) in a specific sense.

**Exercise 1.1** *Let* $A \in \mathbb{R}^{n\times n}$ *is a symmetric positive definite matrix,* $b \in \mathbb{R}^n$*. Prove that the following two problems have a same solution:*

1. *Find* $x \in I\!R^n$ *such that*

$$Ax = b.$$

2. *Find* $x \in I\!R^n$, *such that*

$$J(x) = \min_{y \in I\!R^n} J(y),$$

*where* $J(y) = \frac{1}{2}(Ax, x) - (x, b)$.

Existence and uniqueness?

■ Existence and uniqueness of the solution of (SP) is a direct result of elementary integration (But need new tools in higher dimension).

■ Uniqueness of the solution of (VP)

**Proof**:

- Let (VP) admit two solutions $u_1, u_2 \in V$:

$$(u_1', v') = (f, v) \ \forall v \in V,$$
$$(u_2', v') = (f, v) \ \forall v \in V.$$

- Subtracting these two equations leads to:

$$\int_0^1 (u_1' - u_2')v' dx = 0 \ \forall v \in V.$$

- Choosing $v = u_1 - u_2 \in V$ results in:

$$\int_0^1 (u_1' - u_2')^2 dx = 0,$$

which means that:

$$u_1'(x) - u_2'(x) = 0 \ \forall x \in [0, 1].$$

- It follows that $u_1(x) - u_2(x) =$ constant on $[0,1]$, which together with the boundary condition $u_1(0) = u_2(0) = 0$ gives that:

$$u_1(x) = u_2(x)\ \forall x \in [0,1].$$

## Introduction to Sobolev spaces $L^2, H^1$ and $H_0^1$

- $V$: a linear space

- $(\cdot,\cdot)$: an inner product in $V$, defined as a bilinear mapping $V \times V \to I\!R$, such that

  $1^\circ$ $(u,v) = (v,u)$ for all $u, v \in V$ (symmetry),
  $2^\circ$ $(v,v) \geqslant 0$ for all $v \in V$ (positivity),
  $3^\circ$ $(v,v) = 0$ if and only if $v = 0$.

- $\|\cdot\|_V$: a norm, defined as a mapping $V \to I\!R$, such that

$1^\circ$ $\|v\| \geqslant 0$ for all $v \in V$,

$2^\circ$ $\|cv\| = |c|\|v\|$ for all $c \in I\!R$ and $v \in V$,

$3^\circ$ $\|u + v\| \leqslant \|u\| + \|v\|$ for all $u, v \in V$,

$4^\circ$ $\|u\| = 0$ if and only if $v = 0$.

- $|\cdot|_V$: a seminorm, defined as a mapping satisfying only the first 3 properties in the norm definition.

- Normed Space: $V$ equipped a norm.

- Hilbert Space: $V$ equipped an inner product, and if any Cauchy sequence converges.

- Banach Space: normed space, and if any Cauchy sequence converges.

- Given two normed spaces: $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$. $\mathscr{L}(V, W)$ denotes the space of linear continuous operators from $V$ to $W$, equipped with the norm:

$$\|L\|_{\mathscr{L}(V,W)} = \sup_{v\in V, v\neq 0} \frac{\|Lv\|_W}{\|v\|_V}.$$

- In particular, if $W = R$, $\mathscr{L}(V,W)$ is called dual space of $V$, denoted by $V'$.

- Duality pairing: the bilinear form $\langle\cdot,\cdot\rangle$ from $V' \times V \longrightarrow R$ is defined by $\langle f, v\rangle := f(v)$.

- Schwarz inequality in a Hilbert space $V$: $|(u,v)| \leqslant (u,u)^{1/2}(v,v)^{1/2}, \ \forall u, v \in V$.

## Sobolev space

- $I = (a,b)$.

- $L^p(\Omega) = \left\{v \Big| \int_\Omega |v|^p dx < \infty\right\}, \ 1 \leqslant p < \infty.$

- $L^2(I)$: space of measurable functions whose square is Lebesgue integrable in $I$, endowed with inner product $(u,v) := \int_I uv dx$, and norm $\|v\|_0 := (v,v)^{1/2}$.

* $L^2(I)$ is a Hilbert space.

- $L^\infty(I) = \{v; \sup_{x\in I} v(x) < \infty\}$, equipped with $L^\infty$-norm: $\|v\|_{L^\infty} = \sup_{x\in I} v(x)$.

- $\mathscr{D}(I)$ or $C_0^\infty(I)$: space of infinitely differentiable functions with compact support. [Remark: $\mathscr{D}(I)$ is not a normable space. A meaning of the convergence of a sequence of functions in $\mathscr{D}(I)$ can be found in Adams p20]

- Distribution: defined as a functional in $\mathscr{D}(I)'$; see [Distribution-Carlsson11, p11] for the meaning of a CONTINUOUS functional in $\mathscr{D}(I)$.

- Derivative in the distribution sense: Given a distribution $f$, i.e., $f \in \mathscr{D}(I)'$, define $g \in \mathscr{D}(I)'$ by:

$$\langle g, v\rangle = (-1)\langle f, v'\rangle, \quad \forall v \in \mathscr{D}(I).$$

$g$ is called the first order derivative of $f$, denoted by $f'$.

*In case both* $f$ *and* $g$ *belong to* $L^2(I)$, *the definition becomes*

$$\int_I g(x)\varphi(x)dx = -\int_I f(x)\varphi'(x)dx,\ \forall\varphi \in C_0^\infty(I).$$

$*$ If $f$ is smooth, then $f'$ coincides with the classical one.

**Example 1.1** *Let*

$$f(x) = |x|,\ \forall x \in (-1, 1).$$

*Then*

$$f'(x) = \begin{cases} -1, & \forall x \in (-1, 0), \\ 1, & \forall x \in (0, 1). \end{cases}$$

Second order derivative of $f(x) = |x|, \ \forall x \in (-1, 1)$:

$$
\begin{aligned}
\langle f'', v \rangle = & \quad (-1)\langle f', v' \rangle \\
= & \quad -\int_0^1 v' dx + \int_{-1}^0 v' dx \\
= & \quad 2v(0), \quad \forall v \in \mathscr{D}(I).
\end{aligned}
$$

Thus $f'' = 2\delta_0$, where $\delta_0$ is the Dirac function, defined by

$$\langle \delta_0, v \rangle = v(0), \ \forall v \in \mathscr{D}(R).$$

$*$ The Dirac function is not a $L^p$ function.

**Example 1.2** . *Show that if $f \in \mathscr{D}'(R)$ and $f = 0$, then $f = C$, where $C$ is a constant. [Hint: see distributions_examples.pdf]*

- $H^1(I) = \{v \in L^2(I), v' \in L^2(I)\}$.

- inner product $(u,v)_1 = (u,v) + (u',v')$.

- norm $\|v\|_1 = \sqrt{(v,v)+(v',v')} = (\|v\|_0^2 + \|v'\|_0^2)^{1/2}$.

- semi-norm $|v|_1 = \|v'\|_0$.

- $H^m(I) = \{v^{(i)} \in L^2(I), i = 0, 1, \cdots, m\}$.

- $H_0^1(I) = \{v \in H^1(I), v(0) = v(1) = 0\}$.

- Poincaré inequality: $\|v\|_0 \leqslant c\|v'\|_0,\ \forall v \in H_0^1(I)$.

**Exercise 1.2** *Prove some alternative forms of the Poincaré inequality:*

$$\|v\|_{L^\infty} \leqslant c_1\|v'\|_0,\ \forall v \in \{v \in H^1(I), v(0) = 0\}.$$

$$\|v\|_0 \leqslant c_2\|v'\|_0,\ \forall v \in \{v \in H^1(I), v(0) = 0\}.$$

**Lemma 1.1** *(Lax-Milgram Lemma) Let* $V$ *be a Hilbert space, endowed with the norm* $\|\cdot\|_V$. *Consider the problem:* $\forall f \in L^2(I)$, *find* $u \in V$, *such that*

$$a(u,v) = (f,v), \ \forall v \in V, \tag{1}$$

*where* $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ *is a bilinear form, i.e.,*

$$a(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 a(u_1, v) + \alpha_2 a(u_2, v), \ \forall \alpha_1, \alpha_2 \in \mathbb{R}, u_1, u_2, v \in V.$$

$$a(u, \beta_1 v_1 + \beta_2 v_2) = \beta_1 a(u, v_1) + \beta_2 a(u, v_2), \ \forall \beta_1, \beta_2 \in \mathbb{R}, u, v_1, v_2 \in V.$$

*Furthermore,* $a(\cdot,\cdot)$ *satisfies*

$$\exists \gamma > 0 : |a(u,v)| \leqslant \gamma \|u\|_V \|v\|_V \quad \forall u, v \in V,$$

$$\exists \alpha > 0 : a(v,v) \geqslant \alpha \|v\|_V^2 \quad \forall v \in V.$$

*Then, problem (1) admits a unique solution* $u$, *and* $u$ *satisfies*

$$\|u\|_V \leqslant \frac{1}{\alpha} \sup_{v \in V, v \neq 0} \frac{(f,v)}{\|v\|_V}.$$

**Remark 1.1** *This Lemma remains true for* $\mathcal{F}(v)$ *in place of* $(f, v)$, *where* $\mathcal{F}(v)$ *is a continuous functional from* $V$ *to* $I\!R$*:*

$$|\mathcal{F}(v)| \leqslant c\|v\|_V, \ \forall v \in V.$$

A direct application of Lemma 1.1 to problem (VP) leads to the existence and uniqueness of the solution.

**Exercise 1.3** *Consider the boundary value problem:*

$$\begin{cases} -u''(x) = f(x), \quad x \in (0,1), \\ u(0) = u'(1) = 0, \end{cases} \tag{2}$$

*where* $f$ *is a given continuous function. Let*

$$V = \{v : v \text{ and } v' \text{ are square integrable on } [0,1], \text{ and } v(0) = 0\}.$$

*The corresponding minimization problem of* $(2)$ *reads: Find* $u \in V$, *such that*

$$\mathcal{F}(u) = \inf_{v \in V} \mathcal{F}(v), \tag{3}$$

*where* $\mathcal{F}$ *is defined as:*

$$\mathcal{F}(v) = \frac{1}{2}(v', v') - (f, v), \ \forall v \in V.$$

*The corresponding variational problem of* $(2)$*: Find* $u \in V$, *such that*

$$(u', v') = (f, v), \forall v \in V. \tag{4}$$

*Prove that:*

*1) All three problems* $(2)$, $(3)$, *and* $(4)$ *are equivalent.*

*2) The problem* $(2)$ *admits one unique solution.*

*3) The solution of* $(4)$ *is unique.*

## Galerkin method

Let $V_h \subset V$ being a subspace of $V$. Consider the problem: find $u_h \in V_h$, such that

$$a(u_h, v_h) = (f, v_h), \ \forall v_h \in V_h, \tag{5}$$

where $a(u_h, v_h) = (u_h', v_h')$.

**Theorem 1.1** *Let $u$ and $u_h$ be resp. the solution of (VF) and $(5)$. Then*

$$|u - u_h|_1 \leqslant \inf_{v_h \in V_h} |u - v_h|_1.$$

### Proof

- $(u' - u_h', v_h') = 0, \ \forall v_h \in V_h$.
- $|u - u_h|_1^2 = (u' - u_h', u' - u_h') = (u' - u_h', u' - v_h' + v_h' - u_h'), \ \forall v_h \in V_h$.
- $|u - u_h|_1^2 = (u' - u_h', u' - v_h') + (u' - u_h', v_h' - u_h') = (u' - u_h', u' - v_h'), \ \forall v_h \in V_h$.
- $|u - u_h|_1^2 \leqslant \|u' - u_h'\|_0 \|u' - v_h'\|_0, \ \forall v_h \in V_h$.

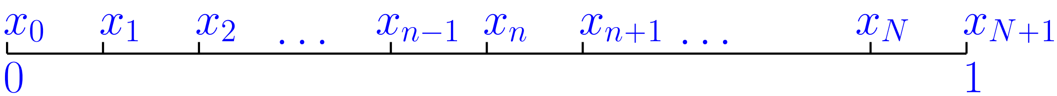
- $|u-u_h|_1^2 \leqslant \|u'-u_h'\|_0\|u'-v_h'\|_0, \ \forall v_h \in V_h$.
- $|u-u_h|_1 \leqslant \inf_{v_h \in V_h} |u-v_h|_1$.

## P1-FEM

Finite element method for the model problem with piecewise linear functions.

Construct a finite-dimensional subspace $V_h \subset V$ as follows:

- Let $\{x_n\}_{n=0}^{N+1}$ be a grid in the interval $I$.
- $I_n = (x_{n-1}, x_n), h_n = x_n - x_{n-1}$.
- $h = \max_{1 \leqslant n \leqslant N+1} h_n$ (the parameter $h$ is a measure of how fine the partition is).



- Let $V_h$ be the space of functions $v_h$ satisfying:
  - $v_h$ is linear on each subinterval $I_n$
  - $v_h$ is continuous on $I$ and
  - $v_h(0) = v_h(1) = 0$.

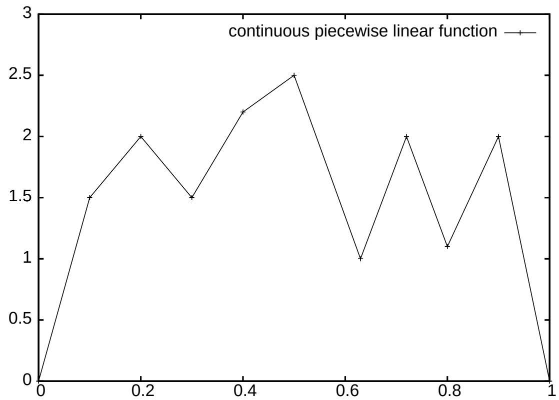
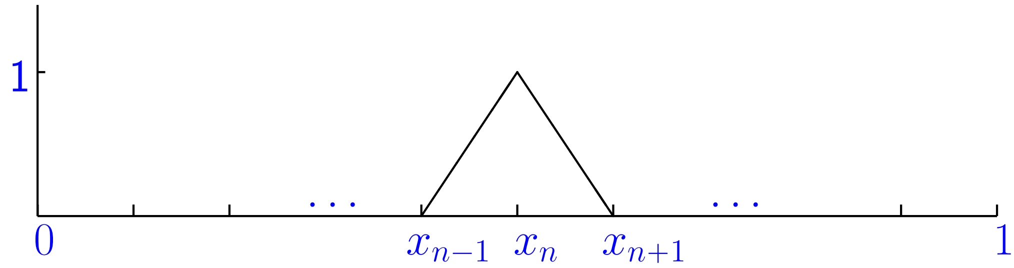



Figure 1: A continuous piecewise linear function.

## Representation of such a function

- Basis functions $\varphi_n \in V_h, n = 1, 2, \cdots, N$, satisfying:

$$\varphi_n(x_m) = \delta_{nm}, \ \forall m = 0, 1, 2, \cdots, N+1.$$




Piecewise linear basis function $\varphi_n$.

- Then

$$V_h = \text{span}\{\varphi_1, \varphi_2, \cdots, \varphi_N\}.$$

- All $v_h \in V_h$ has expression

$$v_h(x) = \sum_{j=1}^{N} v_j \varphi_j(x), \ \forall x \in (0,1),$$

with $v_j = v_h(x_j)$.

- Finite element approximation problem (MP$_h$): Find $u_h \in V_h$, such that

$$\mathcal{F}(u_h) = \min_{v_h \in V_h} \mathcal{F}(v_h). \tag{6}$$

or (VP$_h$): Find $u_h \in V_h$, such that

$$(u_h', v_h') = (f_h, v_h), \ \forall v_h \in V_h. \tag{7}$$

Let $\boldsymbol{v} = (v_1, v_2, \cdots, v_N)^T$, for all $\boldsymbol{v} \in \mathbb{R}^N$, define $J(\boldsymbol{v})$ by $J(\boldsymbol{v}) = \mathcal{F}(v_h)$, where $v_h = \sum_{j=1}^N v_j \varphi_j$. Then, by the definition of $\mathcal{F}$,

$$\begin{aligned} J(\boldsymbol{v}) &= \frac{1}{2}\Big(\sum_{j=1}^N v_j\varphi_j', \sum_{j=1}^N v_j\varphi_j'\Big) - \Big(f, \sum_{j=1}^N v_j\varphi_j\Big) \\ &= \frac{1}{2}\sum_{i,j=1}^N (\varphi_i', \varphi_j') v_i v_j - \sum_{j=1}^N v_j (f, \varphi_j) \\ &= \frac{1}{2}\sum_{i,j=1}^N a_{ij} v_i v_j - \sum_{j=1}^N v_j f_j \\ &= \frac{1}{2}(A\boldsymbol{v}, \boldsymbol{v}) - (\boldsymbol{f}, \boldsymbol{v}), \end{aligned}$$

where

$$A = (a_{ij}), \quad a_{ij} = (\varphi_i', \varphi_j'), \ \forall i, j = 1, 2, \cdots, N,$$

$$\boldsymbol{f} = (f_1, f_2, \cdots, f_N)^T, \quad f_j = (f, \varphi_j), \ \forall j = 1, 2, \cdots, N.$$

$A$: stiffness matrix

Thus, finite dimensional minimization problem (MP$_h$) is equivalent to: Find $\boldsymbol{u} \in I\!R^N$, such that

$$J(\boldsymbol{u}) = \min_{\boldsymbol{v} \in I\!R^N} J(\boldsymbol{v}).$$

Finite dimensional variational problem (VP$_h$) is equivalent to: Find $u_h \in V_h$, such that

$$(u_h', \varphi_j') = (f, \varphi_j), \ j = 1, 2, \cdots, N.$$

or to: Find $\boldsymbol{u} \in I\!R^N$, such that

$$A\boldsymbol{u} = \boldsymbol{f}.$$

Properties of the stiffness matrix $A$

- $A$ is symmetric, $a_{ij} = a_{ji}$, i.e., $(\varphi_i', \varphi_j') = (\varphi_j', \varphi_i'), i, j = 1, 2, \cdots, N$.
- $A$ is sparse (i.e. only a few elements of $A$ are nonzero)

$$(\varphi_j', \varphi_j') = \int_{x_{j-1}}^{x_j} \frac{1}{h_j^2} dx + \int_{x_j}^{x_{j+1}} \frac{1}{h_{j+1}^2} dx = \frac{1}{h_j} + \frac{1}{h_{j+1}}, j = 1, 2, \cdots, N.$$

$$(\varphi_j', \varphi_{j-1}') = (\varphi_{j-1}', \varphi_j') = -\int_{x_{j-1}}^{x_j} \frac{1}{h_j^2} dx = -\frac{1}{h_j}, j = 1, 2, \cdots, N.$$

$$(\varphi_i', \varphi_j') = 0 \text{ if } |i - j| > 1.$$

- $A$ is positive definite. Indeed for $\boldsymbol{v} \in I\!R^N$ we have

$$(A\boldsymbol{v}, \boldsymbol{v}) = \sum_{i,j=1}^N a_{ij} v_i v_j = \sum_{i,j=1}^N (\varphi_i', \varphi_j') v_i v_j = (\sum_{i=1}^N v_i \varphi_i', \sum_{j=1}^N v_j \varphi_j') = (v_h', v_h') \geqslant 0.$$

$$(A\boldsymbol{v},\boldsymbol{v}) = 0 \text{ if and only if } v_j = 0, j = 1, \cdots, N.$$

- $A$ is non-singular, the system $A\boldsymbol{u} = \boldsymbol{f}$ has a unique solution.

Particular case: $h_j = h = \frac{1}{N+1}$

$$A = \frac{1}{h}\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ 0 & \cdots & 0 & 0 & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

## Error estimate

By virtue of the optimal estimate in Theorem 1.1, we have

$$\|u' - u_h'\|_0 \leqslant \inf_{v_h \in V_h} \|u' - v_h'\|_0 \leqslant \|u' - u_I'\|_0, \tag{8}$$

where $u_I$ is the finite element interpolant of $u$ in $V_h$, i.e.,

$$u_I(x_j) = u(x_j), j = 0, 1, 2, \cdots, N+1.$$

Interpolation error:

$$\|u' - u_I'\|_{L^\infty} \leqslant h \max_{x \in I} |u''(x)|, \ \|u - u_I\|_{L^\infty} \leqslant \frac{h^2}{8} \max_{x \in I} |u''(x)|$$
$$\|u' - u_I'\|_0 \leqslant h \max_{x \in I} |u''(x)|, \ \|u - u_I\|_0 \leqslant \frac{h^2}{8} \max_{x \in I} |u''(x)|.$$

A direct application of the above result leads to

$$\|u' - u_h'\|_0 \leqslant h \max_{x \in I} |u''(x)|,$$
$$\|u - u_h\|_0 \leqslant h \max_{x \in I} |u''(x)|,$$
$$\|u - u_h\|_{L^\infty} \leqslant h \max_{x \in I} |u''(x)|.$$

Conclusion: under assumption that $u''$ is bounded on $[0, 1]$, $u_h$ converges to the exact solution $u$ as the maximal length of the subinterval $I_j$ tends to zero.

**Remark 1.2** *This second and third estimates are obtained by using Poincaré inequality, which is not optimal.*

Proof of $\|u - u_I\|_{L^\infty} \leqslant h^2 \max_{x \in I} |u''(x)|$.
By definition, we have

$$u_I(x)|_{I_{j+1}} = u(x_j)\frac{x_{j+1} - x}{h_{j+1}} + u(x_{j+1})\frac{x - x_j}{h_{j+1}}.$$

By using Taylor development,

$$u(x_m) = u(x) + u'(x)(x_m - x) + \frac{1}{2}u''(\xi_m)(x_m - x)^2, \xi_m \in I_{j+1}, m = j, j+1. \quad (9$$

Thus

$$\begin{aligned} u_I(x)|_{I_{j+1}} &= \Big[u(x) + u'(x)(x_j - x) + \frac{1}{2}u''(\xi_j)(x_j - x)^2\Big]\frac{x_{j+1} - x}{h_{j+1}} \\ &\quad + \Big[u(x) + u'(x)(x_{j+1} - x) + \frac{1}{2}u''(\xi_{j+1})(x_{j+1} - x)^2\Big]\frac{x - x_j}{h_{j+1}} \\ &= u(x) + u'(x)\Big[\frac{(x_j - x)(x_{j+1} - x)}{h_{j+1}} + \frac{(x - x_j)(x_{j+1} - x)}{h_{j+1}}\Big] \\ &\quad + \frac{1}{2}u''(\xi_j)(x_j - x)^2\frac{x_{j+1} - x}{h_{j+1}} + \frac{1}{2}u''(\xi_{j+1})(x_{j+1} - x)^2\frac{x - x_j}{h_{j+1}}. \end{aligned}$$

$$
\begin{aligned}
&\|u-u_I\|_{L^\infty(I_{j+1})}\\
= \ & \max_{x\in I_{j+1}} \Big|\frac{1}{2}u''(\xi_j)(x_j-x)^2\frac{x_{j+1}-x}{h_{j+1}}+\frac{1}{2}u''(\xi_{j+1})(x_{j+1}-x)^2\frac{x-x_j}{h_{j+1}}\Big|\\
\leqslant \ & \frac{1}{2}\frac{(x-x_j)(x_{j+1}-x)}{h_{j+1}}((x-x_j)+(x_{j+1}-x))\max_{x\in I_{j+1}}|u''(x)|.\\
\leqslant \ & \frac{h_{j+1}^2}{8}\max_{x\in I_{j+1}}|u''(x)|.
\end{aligned}
$$

(hint: $(x-x_j)(x_{j+1}-x)\leqslant \frac{h_{j+1}^2}{4}, \forall x\in I_{j+1}$)

Proof of $\|u'-u_I'\|_{L^\infty}\leqslant h\max_{x\in I}|u''(x)|$.

Let $x\in I_{j+1}=[x_j,x_{j+1}]$, we prove that

$$
\max_{x\in I_{j+1}}|u'(x)-u_I'(x)|\leqslant h\max_{x\in I}|u''(x)|,\ \forall j=0,1,\cdots,N.
$$

- $u_I'(x)|_{I_{j+1}} = \frac{u(x_{j+1})-u(x_j)}{h_{j+1}}$.
- Using $(9)$ gives

$$u'(x) - u_I'(x) = \frac{1}{2h_{j+1}}[u''(\xi_j)(x_j - x)^2 - u''(\xi_{j+1})(x_{j+1} - x)^2].$$

$$\|u' - u_I'\|_{L^\infty(I_{j+1})} \leqslant h_{j+1} \max_{x \in I_{j+1}} |u''(x)|.$$

$$\|u' - u_I'\|_{L^\infty(I)} \leqslant h\|u''\|_{L^\infty(I)}.$$

**Remark 1.3**

$$\|u' - u_I'\|_0 \leqslant h\|u''\|_{L^\infty(I)}.$$

*In fact, we can prove a slightly better estimate as follows; see [Brenner_Scott_MathematicalTheory-FEM2008.pdf, p8],*

$$\|u' - u_I'\|_0 \leqslant h\|u''\|_0.$$

[Hint:
1) prove $\int_0^1 f(X)^2 dX \leqslant c_1 \int_0^1 f'(X)^2 dX$ and $\int_0^1 f'(X)^2 dX \leqslant c_2 \int_0^1 f''(X)^2 dX$ for all $f \in H^2(I) \cap H_0^1(I)$;
2) make variable change $x = x_{j-1} + X(x_j - x_{j-1})$ to yield $\int_{I_j} \tilde{f}(x)^2 dx \leqslant c_1(x_j - x_{j-1})^2 \int_{I_j} \tilde{f}'(x)^2 dx$ and $\int_{I_j} \tilde{f}'(x)^2 dx \leqslant c_2(x_j - x_{j-1})^2 \int_{I_j} \tilde{f}''(x)^2 dx$ with $\tilde{f}(x) := f(X)$;
3) apply these inequalities to $\tilde{f} = e := u - u_I$. ]

**Remark 1.4** - *The estimate* $(8)$ *can only be used to estimate* $\|u' - u_h'\|_0$ *(not* $\|u - u_h\|_0$*);*

- *A better estimate for* $\|u - u_h\|_0$ *can be obtained by using Aubin-Nitsche trick:*

$$\|u - u_h\|_0 \leqslant h\|u' - u_h'\|_0.$$

## Summary of the $P1$-FEM to BVP

$$\begin{cases} -u''(x) = f(x), \ \forall x \in (0,1), \\ u(0) = u(1) = 0. \end{cases}$$

1) (VP): find $u \in V$, such that

$$a(u,v) = \mathcal{F}(v), \ \forall v \in V,$$

where $V = H_0^1(I), a(u,v) = (u', v'), \mathcal{F}(v) = (f, v)$.

2) (VP$_h$): find $u_h \in V_h$, such that

$$a(u_h, v_h) = \mathcal{F}(v_h), \ \forall v_h \in V_h,$$

where $V_h = \{v_h \in V, v_h$ is linear on each subinterval $I_n\}$.

3) $I = \cup_{n=0}^{N} I_n, I_n = [x_n, x_{n+1}]$.

$$V_h = \mathsf{span}\{\varphi_1, \varphi_2, \cdots, \varphi_N\},$$

where $\varphi_j \in V_h$ such that $\varphi_j(x_n) = \delta_{jn},\ \forall n = 0, 1, \cdots, N+1$.

4) Derive the linear system, investigate the properties of the system matrix.

5) Error analysis

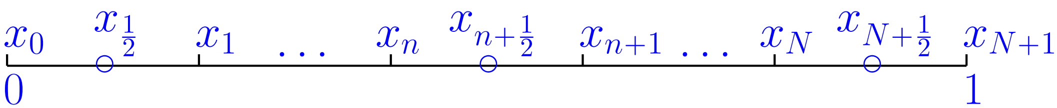
## $P2$-FEM

$$V_h = \{v_h \in C^0(I); v_h|_{I_n} \in \mathbb{P}_2(I_n), \forall n = 0, 1, \cdots, N; v_h(0) = v_h(1) = 0\},$$

where $I_n = [x_n, x_{n+1}]$, $\mathbb{P}_k(I_n)$ is the space of polynomials of degree $\leqslant k$ defined in $I_n$.

Representation of such a piecewise polynomial function

- $Dim(V_h) = 2N + 1$;



- Basis functions $\varphi_n, n = 1, 2, \cdots, N; \varphi_{n+\frac{1}{2}}, n = 0, 2, \cdots, N$, satisfying:

$$\varphi_n(x_m) = \delta_{nm}, \forall m = 0, 1, 2, \cdots, N+1,$$

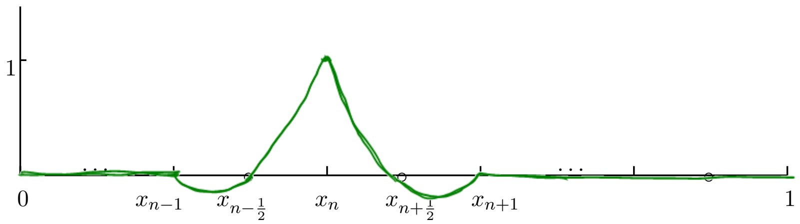
$$\varphi_n(x_{m+\frac{1}{2}}) = 0, \ \forall m = 0, 1, 2, \cdots, N;$$

$$\varphi_{n+\frac{1}{2}}(x_m) = 0, \ \forall m = 0, 1, 2, \cdots, N+1,$$

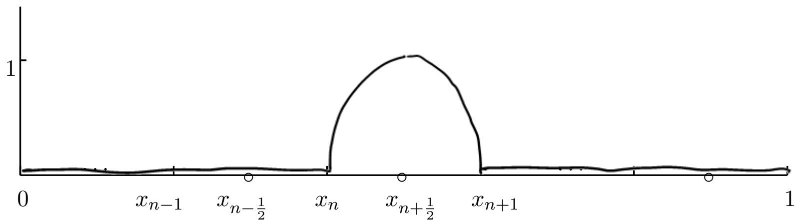
$$\varphi_{n+\frac{1}{2}}(x_{m+\frac{1}{2}}) = \delta_{nm}, \ \forall m = 0, 1, 2, \cdots, N.$$

$$\varphi_n(x) = \begin{cases} \Big(\dfrac{2(x_n - x)}{h_n} - 1\Big)\Big(\dfrac{x_n - x}{h_n} - 1\Big) & x \in I_n \\ \Big(\dfrac{2(x - x_n)}{h_{n+1}} - 1\Big)\Big(\dfrac{x - x_n}{h_{n+1}} - 1\Big) & x \in I_{n+1} \\ 0 & \text{other} \end{cases}$$

$$\varphi_{n+\frac{1}{2}}(x) = \begin{cases} 4\dfrac{x - x_n}{h_{n+1}}\Big(1 - \dfrac{x - x_n}{h_{n+1}}\Big) & x \in I_{n+1} \\ 0 & \text{other} \end{cases}$$




Piecewise $\mathbb{P}_2$ polynomial basis function $\varphi_n$.




Piecewise $\mathbb{P}_2$ polynomial basis function $\varphi_{n+\frac{1}{2}}$.

Figure 2: Plot of the piecewise $\mathbb{P}_2$ basis functions.

- Then

$$V_h = \text{span}\{\varphi_{\frac{1}{2}}, \varphi_1, \varphi_{1+\frac{1}{2}}, \varphi_2, \cdots, \varphi_N, \varphi_{N+\frac{1}{2}}\}.$$

- All $v_h \in V_h$ has expression

$$v_h(x) = \sum_{j=1}^{N} v_j \varphi_j(x) + \sum_{j=0}^{N} v_{j+\frac{1}{2}} \varphi_{j+\frac{1}{2}}(x), \ \forall x \in (0,1),$$

with $v_j = v_h(x_j), v_{j+\frac{1}{2}} = v_h(x_{j+\frac{1}{2}})$.

- Error estimate:

$$\|u' - u_h'\|_0 = O(h^2), \quad \|u - u_h\|_0 = O(h^3)$$

compared to $P_1$-FEM based on the same grid points:

$$\|u' - u_h'\|_0 = O(h), \quad \|u - u_h\|_0 = O(h^2).$$

## $P3$-FEM

$$V_h = \{v_h \in C^0(I); v_h|_{I_n} \in \mathbb{P}_3(I_n), \forall n = 0, 1, \cdots, N; v_h(0) = v_h(1) = 0\}.$$

Similar analysis applies!

## Other homogeneous boundary conditions

- Neumann condition

$$\begin{cases} u - u''(x) = f(x), \ \forall x \in (0,1), \\ u'(0) = u'(1) = 0. \end{cases}$$

(VP): $V = H^1(I), a(u,v) = (u,v) + (u',v'), \mathcal{F}(v) = (f,v)$.

$$V_h = \text{span}\{\varphi_0, \varphi_1, \cdots, \varphi_{N+1}\}.$$

- Mixed condition

$$\begin{cases} -u''(x) = f(x), \ \forall x \in (0,1), \\ u(0) = u'(1) = 0. \end{cases}$$

(VP): $V=\{\boldsymbol{v}\in H^1(I), v(0)=0\}, a(u,v)=(u',v'), \mathcal{F}(v)=(f,v)$.

$$V_h=\text{span}\{\varphi_1,\varphi_2,\cdots,\varphi_{N+1}\}.$$

- Robin condition

$$\begin{cases} -u''(x)=f(x),\ \forall x\in(0,1), \\ u(0)-u'(0)=0, \\ u(1)+u'(1)=0. \end{cases}$$

(VP): $V=H^1(I), a(u,v)=(u',v')+u(0)v(0)+u(1)v(1), \mathcal{F}(v)=(f,v)$.

$$V_h=\text{span}\{\varphi_0,\varphi_1,\cdots,\varphi_{N+1}\}.$$

## Non-homogeneous boundary conditions

- Dirichlet condition

$$u(0)=\alpha,\ u(1)=\beta.$$

Homogenization

$$u = \bar{u} + u^*,\ u^*(0) = \alpha,\ u^*(1) = \beta.$$

- Neumann condition

$$u'(0) = \alpha,\ u'(1) = \beta.$$

(VP): $V = H^1(I), a(u,v) = (u,v) + (u',v'), \mathcal{F}(v) = (f,v) + \beta v(1) - \alpha v(0)$.

- Mixed condition

$$u(0) = \alpha,\ u'(1) = \beta.$$

Homogenization

$$u = \bar{u} + u^*,\ u^*(0) = \alpha.$$

- Robin condition

$$\begin{cases} -u''(x) = f(x), \ \forall x \in (0,1), \\ u(0) - u'(0) = \alpha, \\ u(1) + u'(1) = \beta. \end{cases}$$

(VP): $V = H^1(I), a(u,v) = (u', v') + u(0)v(0) + u(1)v(1), \mathcal{F}(v) = (f, v) + \beta v(1) + \alpha v(0)$.

**Exercise 2.1** *Consider the following problems.*

- *Derive the variational formulation;*

- *Establish the existence and uniqueness;*

- *Construct a $P1-$FEM method.*

*1) Helmholtz Dirichlet problem:* $\alpha \geqslant 0$

$$\begin{cases} \alpha u - u''(x) = f(x), \quad x \in (0,1), \\ u(0) = u(1) = 0. \end{cases}$$

*2) Helmholtz Neumann problem:* $\alpha > 0$

$$\begin{cases} \alpha u - u''(x) = f(x), \quad x \in (0,1), \\ u'(0) = u'(1) = \beta. \end{cases}$$

*3) Helmholtz mixed problem:* $\alpha \geqslant 0$

$$\begin{cases} \alpha u - u''(x) = f(x), \quad x \in (0,1), \\ u(0) = 0, u'(1) = \beta. \end{cases}$$

**Exercise 2.2** *Consider*

$$\begin{cases} u(x) + u'(x) - u''(x) = f(x), \quad x \in I, \\ u(0) = u(1) = 0. \end{cases}$$

$1^o$ *Prove the problem is equivalent to: find* $u \in H_0^1(I)$, *such that*

$$(u,v) + (u',v) + (u',v') = (f,v), \ \forall v \in H_0^1(I). \tag{10}$$

$2^o$ *Prove problem* $(10)$ *admits a unique solution.*

$3^o$ *Let*

$$J(v) = \frac{1}{2}[(v,v) + (v',v) + (v',v')] - (f,v), \ \forall v \in H_0^1(I).$$

*Question: is problem* $(10)$ *equivalent to: find* $u \in H_0^1(I)$*, such that*

$$J(u) = \min_{v \in H_0^1(I)} J(v).$$

$4^o$ *If the boundary condition is replaced by* $u(0) = u'(1) = 0$*, what is the situation?*

**Exercise 2.3** *(Numerical experiments) Let* $\alpha \geqslant 0$*,* $k$ *is an integer. Solve numerically the problem by* $\mathbb{P}_1$*-FE and* $\mathbb{P}_2$*-FE methods:*

$$\begin{cases} \alpha u - u'' = f(x), \ x \in (0,1), \\ u(0) = a, \\ u'(1) = 2\pi k. \end{cases}$$

*Take* $f(x) = a\alpha + (\alpha + 4\pi^2 k^2)\sin(2\pi kx)$ *such that* $u(x) = a + \sin(2\pi kx)$.

*1) Investigate the convergence rate with respect to the mesh size* $h$*;*

*2) Investigate the impact of the parameter* $\alpha, a$ *and* $k$ *on the accuracy.*

[Outline of the report:

Title: Numerical investigations of finite element methods for elliptic equations

Abstract: This paper aims to numerically investigate the accuracy of finite element methods for elliptic equations. Precisely, we consider a Dirichlet problem of an elliptic equation, and propose P1-FEM and P2-FEM for this problem. The theoretical convergence order of the proposed methods is proved. A series of numerical examples are provided to verify the theoretical results.

Section 1. Problem and numerical methods

Section 2. Implementation and numerical analysis

Section 3. Numerical experiments

Section 4. Conclusion ]

**Exercise 2.4** *We consider the problem*

$$\begin{cases} -(\alpha u')'(x) + (\beta u')(x) + (\gamma u)(x) = f(x), \quad x \in (0,1) \\ u(0) = u(1) = 0, \end{cases} \tag{11}$$

*where* $\alpha, \beta$, *and* $\gamma$ *are continuous functions on* $[0,1]$ *with* $\alpha(x) \geqslant \alpha_0 > 0$ *for all* $x \in [0,1]$.

*1) Give the weak form of the problem* $(11)$.

*2) Prove the weak problem admits a unique solution under the following assumption*

*a.* $\beta(x) = 0, \gamma \geqslant 0$ *for all* $x \in [0,1]$;

*or*

*b.* $-\frac{1}{2}\beta' + \gamma \geqslant 0$ *for all* $x \in [0,1]$.

*or*

*c. see [Brezis p224].*

*3) Propose a P1-FEM for the numerical solution of* $(11)$.

*4) Carry out an error analysis.*

**Exercise 2.5** *Advection-Diffusion Equations:*

$$\begin{cases} -\varepsilon u''(x) + \beta u'(x) = 0, & x \in (0,1), \\ u(0) = 0, u(1) = 1, \end{cases} \tag{12}$$

*where* $\varepsilon$ *and* $\beta$ *are two positive constants such that* $\varepsilon/\beta << 1$. *Define the global Péclet number as*

$$Pe_{gl} = \frac{|\beta| L}{2\varepsilon},$$

*where* $L$ *is the size of the domain (equal to 1 in our case). The exact solution:*

$$u(x) = \frac{e^{\beta x/\varepsilon} - 1}{e^{\beta/\varepsilon} - 1}.$$

Numerically solve this problem by using P1-FEM.

**Exercise 2.6** *(Application 1) Lubrication of a Slide:*

$$\begin{cases} -\left(\frac{s^3}{6\mu}p'\right)'(x) = -(Us)', \quad x \in (0, L), \\ p(0) = p(L) = 0, \end{cases} \tag{13}$$

*where* $L = 1, s(x) = 1 - \frac{3}{2}x + \frac{9}{8}x^2, \mu = 1$.
*Solve this problem by using P1-FEM and P2-SEM.*

**Exercise 2.7** *(Application 2) Vertical Distribution of Spore Concentration over Wide Regions:*

$$\begin{cases} -\nu u''(x) + \beta u'(x) = 0, \quad x \in (0, H), \\ u(0) = u_0, \quad -\nu u'(H) + \beta u(H) = 0. \end{cases} \tag{14}$$

*where* $H$ *is a fixed height at which we assume a vanishing Neumann condition. Realistic values of the coefficients are* $\nu = 10m^2s^{-1}$ *and* $\beta = -0.03ms^{-1}$. *As for*

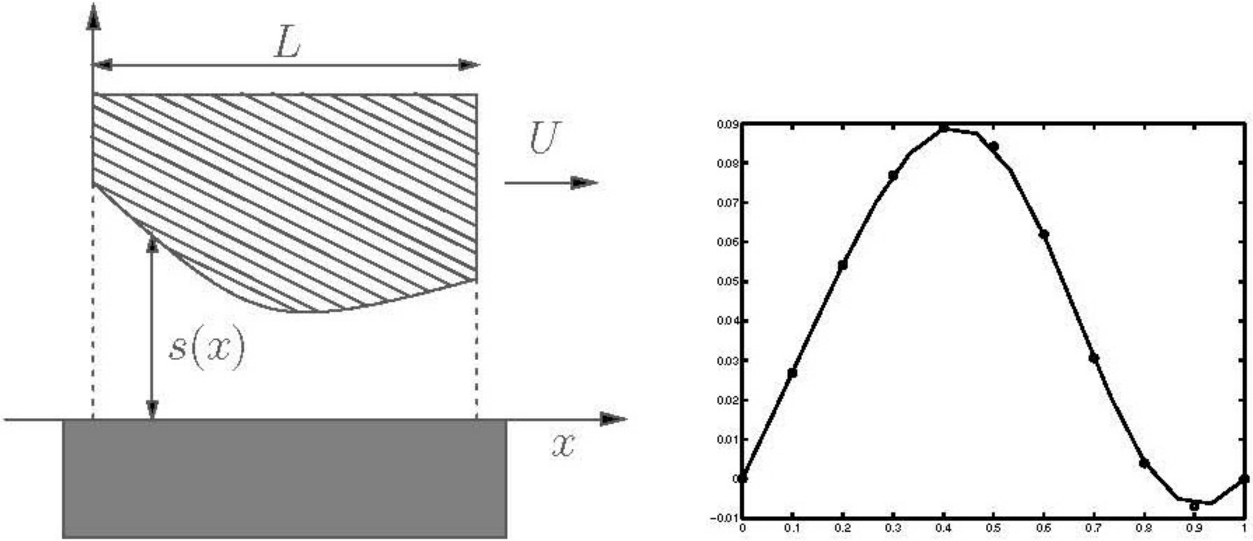



Figure 3: Left: geometrical parameters of the slider; right: pressure on a converging-diverging slider. The solid line denotes the solution obtained used linear finite elements, while the dashed line denotes the solution obtained using quadratic finite elements.

$u_0$, *we take a reference concentration of 1 pollen grain per* $m^3$, *while the height* $H$ *is set equal to* $10km$. *The global Péclet number is therefore* $Pe_{gl} = 15$. *Find the numerical solution of this problem by using P1-FEM and P2-SEM.*

# 2D model problem

Consider the Poisson problem (SP):

$$\begin{cases} -\triangle u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \\ u|_\Gamma = 0, \end{cases}$$

where

- $\Omega \subset I\!R^2, \mathbf{x} = (x, y)$;
- $\Gamma$ is the boundary of $\Omega$, denoted also $\partial\Omega$;
- $\triangle u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$;

$*$ (SP) models the displacement of an elastic membrane fixed at the boundary under a load $f$.

The Minimization Problem (MP): Find $u \in V$, such that $\mathcal{F}(u) \leqslant \mathcal{F}(v), \ \forall v \in V$, where $\mathcal{F}$ is the linear functional $V \to R$, defined as:

$$\mathcal{F}(v) = \frac{1}{2}(\nabla v, \nabla v) - (f, v), \ \forall v \in V,$$

where

- $V = \{v : v \in H^1(\Omega), v|_\Gamma = 0\}$;
- $(v, w) = \int_\Omega v(\mathbf{x})w(\mathbf{x})d\mathbf{x}$, for all scalar function $v, w \in V$;
- $(\nabla v, \nabla w) = \int_\Omega \left(\frac{\partial v}{\partial x}\frac{\partial w}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial w}{\partial y}\right)d\mathbf{x}$;

The Variational Problem (VP): Find $u \in V$, such that

$$(\nabla u, \nabla v) = (f, v), \forall v \in V.$$

## Relationship between (SP), (MP), and (VP)

## The solution of (SP) is also a solution of (VP)

**Proof**:

- Multiplying the equation $-\triangle u = f$ by an arbitrary function $v \in V$;
- Integrating over $\Omega$ which gives:

$$(-\triangle u, v) = (f, v), \ \forall v \in V;$$

- Applying the integration by parts in the left-hand side and using the fact that $v = 0$ on $\Gamma$ to get:

$$-(\triangle u, v) = (\nabla u, \nabla v) - \int_\Gamma \frac{\partial u}{\partial \mathbf{n}} v d\sigma = (\nabla u, \nabla v);$$

- Finally, we conclude that: $(\nabla u, \nabla v) = (f, v)$ for all $v \in V$ , i.e. $u$ is a solution of (VP).

## The solution of (VP) is also a solution of (SP)

**Proof**:

- Suppose we have

$$(\nabla u, \nabla v) = (f, v), \ \forall v \in V;$$

- Integration by parts in the left-hand side gives:

$$(-\triangle u, v) = (f, v), \ \forall v \in V.$$

Thus

$$(-\triangle u, v) = (f, v), \ \forall v \in C_0^\infty(\Omega).$$

- Basic Lemma leads to

$$-\triangle u = f, \ \forall x \in \Omega.$$

## Equivalence between (VP) and (MP)

**Proof**: replace $\cdot'$ by $\nabla\cdot$ in 1D case.

Mixed problem (SP)

$$\begin{cases} -\triangle u(\mathbf{x}) = f(\mathbf{x}), & \mathbf{x} \in \Omega, \\ u|_{\Gamma_D} = 0, \\ \frac{\partial u}{\partial \mathbf{n}}|_{\Gamma_N} = g, \end{cases}$$

where $g \in L^2(\Gamma_N)$, $\Gamma_D \subset \partial\Omega, \Gamma_N \subset \partial\Omega, \Gamma_D \cup \Gamma_N = \partial\Omega$.

The Variational Problem (VP): Find $u \in V$, such that

$$a(u,v) = \mathcal{F}(v), \forall v \in V,$$

where

- $V = \{v \in H^1(\Omega), v|_{\Gamma_D} = 0\}$;
- $a(u, v) = (\nabla u, \nabla v)$;
- $\mathcal{F}(v) = (f, v) + \int_{\Gamma_N} g v d\sigma$.

**Theorem** For $g \in L^2(\Gamma_N)$, $\Gamma_D \neq \varnothing$, the problem (VP) admits a unique solution. Moreover, the solution of (VP) is also a solution of (SP). Finally, the solution $u$ satisfies

$$\|u\|_{1,\Omega} \leqslant c(\|f\|_{0,\Omega} + \|g\|_{0,\Gamma_N}).$$

**Proof** Applying Lax-Milgram Lemma.

**Exercise 3.1** *Let $\Omega = (a, b)^2, f \in L^2(\Omega)$. Consider the Dirichlet elliptic problem*

$$\begin{cases} -\triangle u(\mathbf{x}) = f(\mathbf{x}), & \mathbf{x} \in \Omega, \\ u|_\Gamma = 0. \end{cases}$$

*1. Prove the following Poincaré inequality holds: there exists a constant $c$, depending only on $a$ and $b$, such that*

$$\|v\|_1 \leqslant c|v|_1, \ \forall v \in H_0^1(\Omega).$$

*2. Prove that the Dirichlet elliptic problem admits a unique weak solution in $H_0^1(\Omega)$, and the solution $u$ satisfies*

$$\|u\|_1 \leqslant c\|f\|_0,$$

*where $c$ is a constant.*

## Triangulation

- $\Omega \subset I\!R^d, d = 2, 3$: polygonal domain
- $\mathcal{T}_h$ is a set of polyhedron
- $\bar{\Omega} = \cup_{K \in \mathcal{T}_h}$

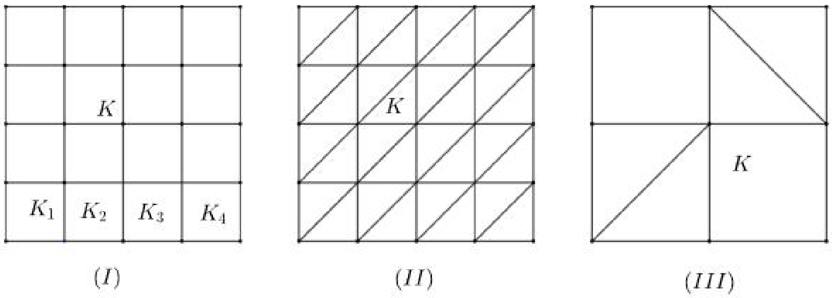
- each $K$ is a polyhedron with $\mathring{K} \neq \varnothing$

- if $F = K_1 \cap K_2 \neq \varnothing$ ($K_1$ and $K_2$ are distinct elements of $\mathcal{T}_h$), then $F$ is a common face, side or vertex of $K_1$ and $K_2$

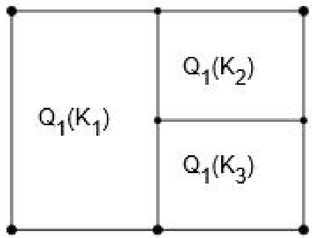
- diam$K \leqslant h, \ \forall K \in \mathcal{T}_h$

$\mathcal{T}_h$ is called a triangulation of $\bar{\Omega}$.

Allowed partition of $\Omega$:



Not allowed partition




## Piecewise polynomial spaces

$$X_h^k = \{v_h \in C^0(\bar{\Omega}), v_h|_K \in \mathcal{P}_k(K), \forall K \in \mathcal{T}_h\}$$

in case $K$ is a triangle, or

$$X_h^k = \{v_h \in C^0(\bar{\Omega}), v_h|_K \in \mathcal{Q}_k(K), \forall K \in \mathcal{T}_h\}$$

in case $K$ is a rectangle.

- $\mathcal{P}_k(K) = \{\sum_{\alpha_1+\cdots+\alpha_d=0}^{k} c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}, \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_d), \alpha_i \geqslant 0\}$

- $\mathcal{Q}_k(K) = \{\sum_{\alpha_1=0,\cdots,\alpha_d=0}^{k} c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}, \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_d), \alpha_i \geqslant 0\}$

- Dim$(\mathcal{P}_k) = \begin{pmatrix} d+k \\ k \end{pmatrix}$

- Dim$(\mathcal{Q}_k) = (k+1)^d$

- $X_h^k \subset H^1(\Omega)$

## Degrees of freedom and shape functions

Construct a basis for $X_h^k$

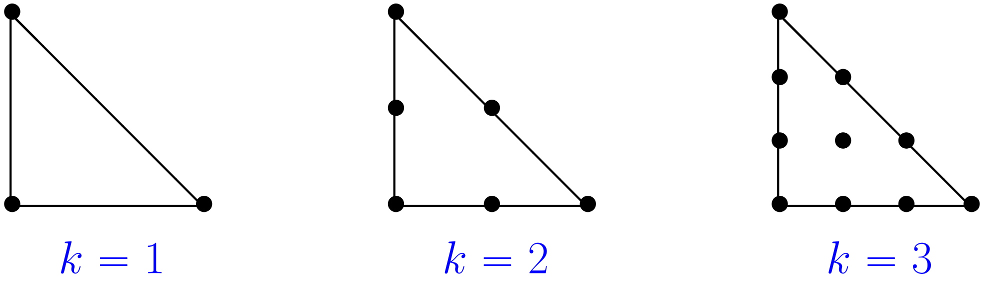
1. Triangular FE

- $d = 2$

- $k=1$: to identify $v_h|_K$ with $k=1$, the simplest choice is values at the vertices of each $K$.

- $k=2$: local dimension is 6

- $k=3$: local dimension is 10




- $d=3$

  easy to generalize.

- Definition: Finite element for a $K$ in $\mathcal{T}_h$.

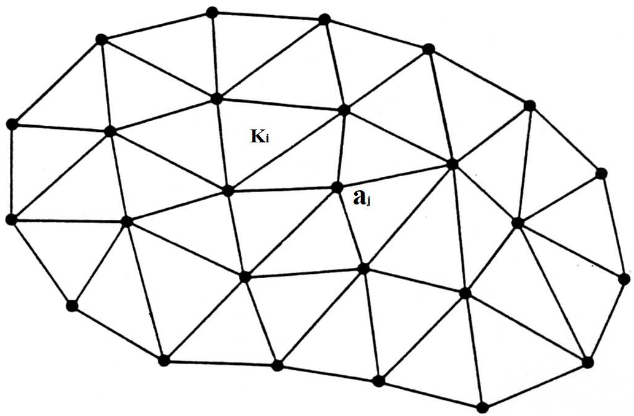



Figure 4: Triangular finite elements.

Let $P(K)$ be a function space, $\sum_K$ is a point set such that the function in $P(K)$ can be uniquely determined by the values at $\sum_K$ (called uni-solved set), then $(K, P(K), \sum_K)$ is called a finite element.

Example of finite elements:
(1) $K$ is a triangle, $P(K) \hat{=} P_1(K) = span\{1, x, y\}, \sum_K = \{a_j | j = 1, 2, 3\}$, with $\{a_j\}$ being the three vertexes.
(2) $K$ is a triangle, $P(K) \hat{=} P_1(K), \sum_K = \{b_j | j = 1, 2, 3\}$, with $\{b_j\}$ being the three midpoints.

Let $N_h$ is the number of the global set of nodes in $\Omega$, $a_j$, the basis functions are all functions $\varphi_i \in X_h^k$, such that

$$\varphi_i(a_j) = \delta_{ij}, \ \forall i, j = 1, \cdots, N_h.$$

$*$ basis function $\varphi_i$ is often called shape function.

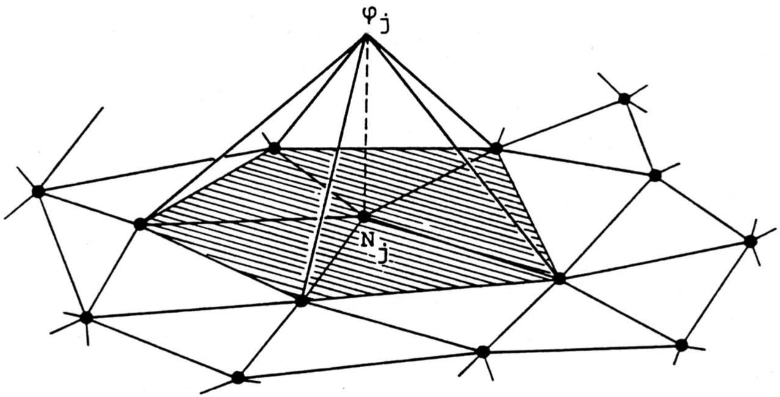
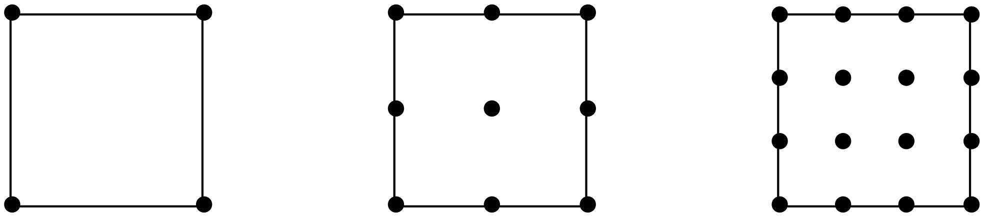



Figure 5: Shape function based on the triangular FE.

## 2. Parallelepipedal FE



$k = 1$ $k = 2$ $k = 3$

Examples:

- if $K$ is a rectangle, $P(K) \hat{=} Q_1(K) = span\{1, x, y, xy\}$, $\sum_K$ = {mid-points of four sides}. Then $(K, P(K), \sum_K)$ is not a finite element (prove it).

- if $K$ is a rectangle, $P(K) \hat{=} Q_1^T(K) = span\{1, x, y, x^2 - y^2\}$, $\sum_K$ = {mid-points of four sides}. Then $(K, P(K), \sum_k)$ is a finite element (rotated element) (prove it).

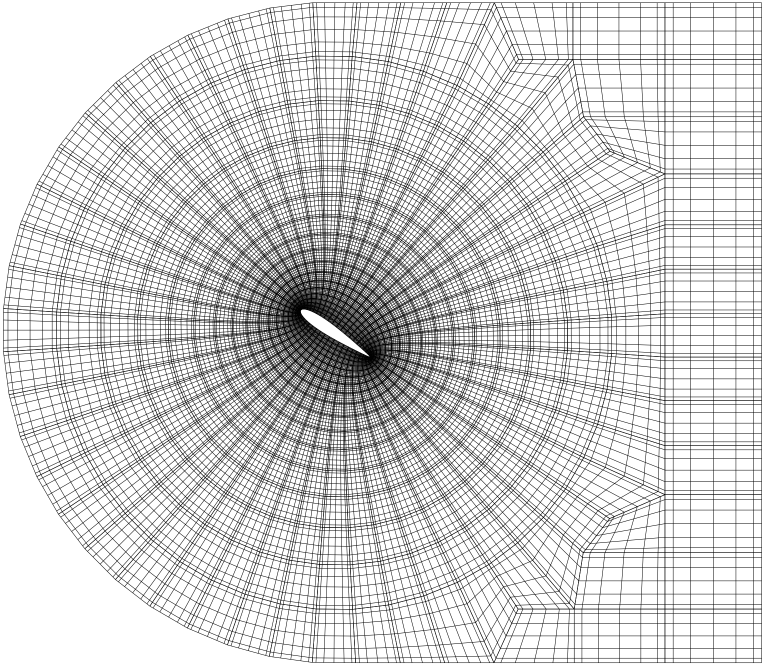


Figure 6: Rectangular FE for an airfoil flow.

## Interpolation operator and error analysis

Definition: Let $h_K = \text{diam}(K)$, $\rho_K = \sup\{\text{diam}(B); B \text{ is a ball in } K\}$. A triangulation $\mathcal{T}_h$ is said regular if there exists $\sigma \geqslant 1$, such that

$$\max_{K \in \mathcal{T}_h} \frac{h_K}{\rho_K} \leqslant \sigma, \ \forall h > 0.$$

$v \in C^0(\bar{\Omega})$, define its interpolant $\pi_h^k v = \sum_{i=1}^{N_h} v(a_i)\varphi_i$.

**Theorem** Let $\mathcal{T}_h$ is regular, $l = \min(k, s-1) \geqslant 1$. Then there exists a constant $c$ independent of $h$, such that

$$|v - \pi_h^k v|_{m,\Omega} \leqslant c h^{l+1-m} |v|_{l+1,\Omega}, \ \forall v \in H^s(\Omega).$$

# Implementation

(VP): Let $\Omega = (0,1)^2$. Find $u \in H_0^1(\Omega)$, such that

$$a(u,v) = \mathcal{F}(v),\ \forall v \in H_0^1(\Omega),$$

where

$$a(u,v) = (\nabla u, \nabla v), \mathcal{F}(v) = (f,v).$$

$\mathcal{Q}_1$-FEM: Find $u_h \in V_h = X_h^1 \cap H_0^1(\Omega)$, such that

$$a(u_h,v_h) = \mathcal{F}(v_h),\ \forall v_h \in V_h.$$

- Rectangular mesh

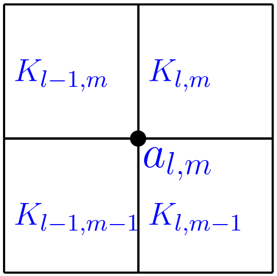
- Nodes are denoted by $a_{l,m}, l,m = 0,1,\cdots,M+1$

- $\varphi_{l,m}$ is the basis function associated to $a_{l,m}$ such that

$$\varphi_{l,m}(a_{p,q}) = \delta_{lp}\delta_{mq},\ \forall p,q = 0,1,\cdots,M+1.$$

- For $l, m = 1, 2, \cdots, M$,

$$
\varphi_{l,m}(x,y) = \begin{cases} \dfrac{x - x_{l-1}}{h} \dfrac{y - y_{m-1}}{h}, & (x,y) \in K_{l-1,m-1} \\ \dfrac{x_{l+1} - x}{h} \dfrac{y - y_{m-1}}{h}, & (x,y) \in K_{l,m-1} \\ \dfrac{x_{l+1} - x}{h} \dfrac{y_{m+1} - y}{h}, & (x,y) \in K_{l,m} \\ \dfrac{x - x_{l-1}}{h} \dfrac{y_{m+1} - y}{h}, & (x,y) \in K_{l-1,m} \\ 0, & others \end{cases}
$$




$\Omega_{l,m}$

- Let $u_h = \sum_{l,m=1}^{M} u_{l,m}\varphi_{l,m}$, $f_{l,m} = \int_\Omega f\varphi_{l,m}dx$
- Straightforward calculation gives

$$\begin{cases} 3u_{l,m} - \frac{1}{3}\sum_{p,q=-1}^{1} u_{l+p,m+q} = f_{l,m}, \ 1 \leqslant l, m \leqslant M \\ u_{l,0} = u_{l,M+1} = 0. \ 0 \leqslant l \leqslant M+1 \\ u_{0,m} = u_{M+1,m} = 0, \ 0 \leqslant m \leqslant M+1 \end{cases}$$

(9-point schema)

Indeed,

$$\begin{aligned} a(\varphi_{l,m}, \varphi_{l,m}) &= \int_{\Omega_{l,m}} \Big[ \Big(\frac{\partial \varphi_{l,m}}{\partial x}\Big)^2 + \Big(\frac{\partial \varphi_{l,m}}{\partial y}\Big)^2 \Big] dxdy \\ &= 4\int_{\Omega_{l-1,m-1}} \Big[ \frac{1}{h^2}\Big(\frac{y-y_{m-1}}{h}\Big)^2 + \frac{1}{h^2}\Big(\frac{x-x_{l-1}}{h}\Big)^2 \Big] dxdy \\ &= \frac{8}{3}. \end{aligned}$$

Similarly, we have

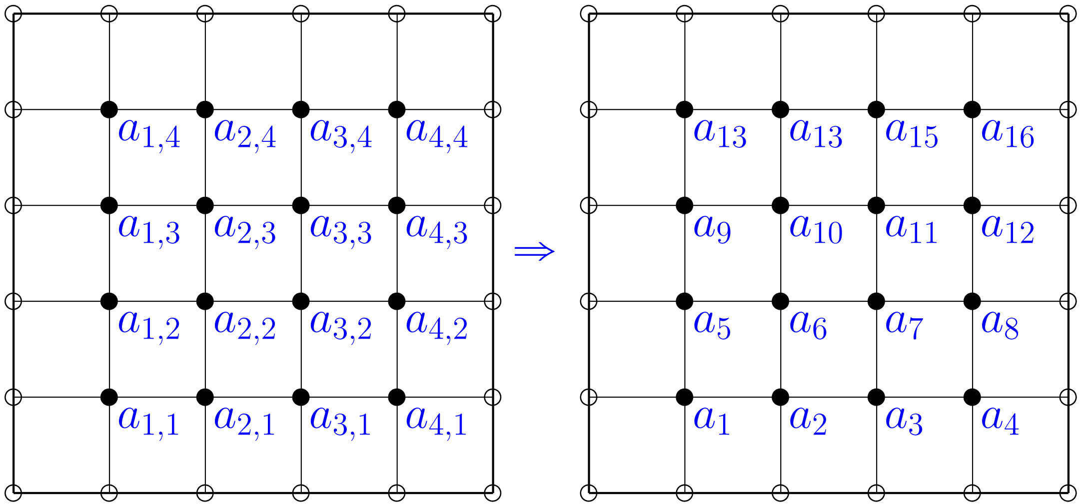
$$
\begin{aligned}
& a(\varphi_{l-1,m-1}, \varphi_{l,m}) \\
= & -\int_{\Omega_{l-1,m-1}} \Big[\frac{1}{h^2}\Big(\frac{y-y_{m-1}}{h}\ \frac{y_m - y}{h}\Big) - \frac{1}{h^2}\Big(\frac{x-x_{l-1}}{h}\ \frac{x-x_{l-1}}{h}\ \frac{x_l - x}{h}\Big)\Big] dxdy \\
= & -2\int_0^1 (Y - Y^2) dY \\
= & -\frac{1}{3}.
\end{aligned}
$$

$$
a(\varphi_{l,m-1}, \varphi_{l,m}) = -\frac{1}{3}.
$$

Matrix structure: an example for $M = 4$.

- Numbering the nodes by

$$
(l, m) \rightarrow i = l + M(m-1)
$$




- Basis functions $\varphi_i, 1 \leqslant i \leqslant M^2$, such that

$$\varphi_i(a_j) = \delta_{ij}, \ 1 \leqslant i, j \leqslant M^2.$$

- System matrix $A = (a_{ij})_{i,j=1}^{M}$, with $a_{ij} = a(\varphi_j, \varphi_i)$:

$$
\begin{bmatrix}
* & * & & & * & * & & & & & & & & & & \\
* & * & * & & * & * & * & & & & & & & & & \\
& * & * & * & & * & * & * & & & & & & & & \\
& & * & * & & & * & * & & & & & & & & \\
* & * & & & * & * & & & * & * & & & & & & \\
* & * & * & & * & * & * & & * & * & * & & & & & \\
& * & * & * & & * & * & * & & * & * & * & & & & \\
& & * & * & & & * & * & & & * & * & & & & \\
& & & & * & * & & & * & * & & & * & * & & \\
& & & & * & * & * & & * & * & * & & * & * & * & \\
& & & & & * & * & * & & * & * & * & & * & * & * \\
& & & & & & * & * & & & * & * & & & * & * \\
& & & & & & & & * & * & & & * & * & & \\
& & & & & & & & * & * & * & & * & * & * & \\
& & & & & & & & & * & * & * & & * & * & * \\
& & & & & & & & & & * & * & & & * & *
\end{bmatrix}
$$

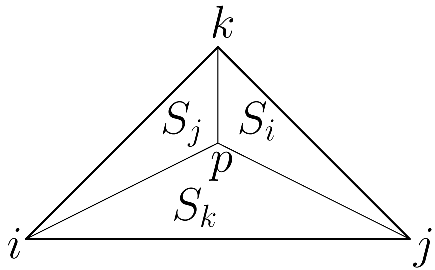
(3-diagonal by block, 3-diagonal each block)

## Barycentric coordinates

Express a piecewise polynomial by using the barycentric coordinates, defined by the vertices of a simplex (a triangle, tetrahedron, etc).

- $\triangle(i,j,k) = \triangle(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$
- $\forall \mathbf{x}_p \in \triangle(i,j,k)$,

let $S_i, S_j, S_k$ be resp. the area of $\triangle(j,k,p), \triangle(k,i,p), \triangle(i,j,p)$




- Then $S_i + S_j + S_k = S, S =$ area of $\triangle(i,j,k)$
- Let $L_i = S_i/S, L_j = S_j/S, L_k = S_k/S$, then

$$L_i + L_j + L_k = 1,\ L_i \geqslant 0,\ L_j \geqslant 0,\ L_k \geqslant 0,$$

$$
\begin{aligned}
p &\longleftrightarrow \{L_i, L_j, L_k\} \\
i &\longleftrightarrow \{1, 0, 0\} \\
j &\longleftrightarrow \{0, 1, 0\} \\
k &\longleftrightarrow \{0, 0, 1\}
\end{aligned}
$$

Relationship between descartes $(x, y)$ and barycentric coordinates $(L_i, L_j, L_k)$:

$$
2S = \begin{vmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{vmatrix}, \qquad 2S_i = \begin{vmatrix} 1 & x & y \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{vmatrix}, \qquad \cdots
$$

$$
\begin{cases}
L_i = \frac{1}{2S}\big[(x_j y_k - x_k y_j) + (y_j - y_k)x + (x_k - x_j)y\big], \\
L_j = \frac{1}{2S}\big[(x_k y_i - x_i y_k) + (y_k - y_i)x + (x_i - x_k)y\big], \\
L_k = \frac{1}{2S}\big[(x_i y_j - x_j y_i) + (y_i - y_j)x + (x_j - x_i)y\big].
\end{cases}
$$

Inversely

$$\begin{cases} x = x_i L_i + x_j L_j + x_k L_k \\ y = y_i L_i + y_j L_j + y_k L_k \end{cases}$$

Derivatives

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial L_i}\frac{\partial L_i}{\partial x} + \frac{\partial}{\partial L_j}\frac{\partial L_j}{\partial x} + \frac{\partial}{\partial L_k}\frac{\partial L_k}{\partial x}$$

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial L_i}\frac{\partial L_i}{\partial y} + \frac{\partial}{\partial L_j}\frac{\partial L_j}{\partial y} + \frac{\partial}{\partial L_k}\frac{\partial L_k}{\partial y}.$$

$\mathcal{P}_1$-FEM space $X_h^1$: $\forall v \in X_h^1$, if $K = \triangle(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$, $v(\mathbf{x}_i) = v_i, v(\mathbf{x}_j) = v_j, v(\mathbf{x}_k) = v_k$, then

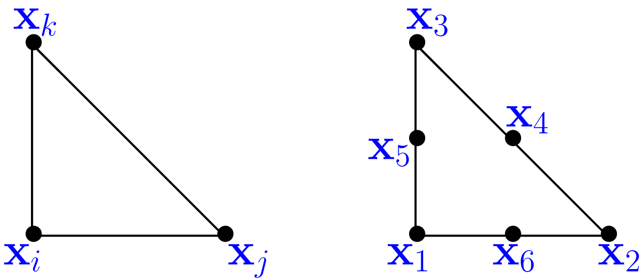
$$v(\mathbf{x})|_K = v_i L_i + v_j L_j + v_k L_k.$$

Similarly, for $v \in X_h^2$, if $K = \triangle(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, $\mathbf{x}_4, \mathbf{x}_5$, and $\mathbf{x}_6$ are the midpoints of the sides, $v(\mathbf{x}_i) = v_i, i = 1, 2, \cdots, 6$, then

$$v(\mathbf{x})|_K = \sum_{i=1}^{3} \left[ v_i L_i (2L_i - 1) + 4 v_{i+3} L_{i+1} L_{i+2} \right]$$

with

$$L_4 = L_1, L_5 = L_2, L_6 = L_3.$$




$\mathcal{P}_1$-element $\mathcal{P}_2$-element

**Exercise 3.2** *Let* $K = \triangle(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ *be a triangle,* $\mathbf{x}_4, \mathbf{x}_5$*, and* $\mathbf{x}_6$ *are the midpoints of the sides,* $p_2(\mathbf{x})$ *is a polynomial of degree* $2$*, such that* $v(\mathbf{x}_i) = v_i, i = 1, 2, \cdots, 6$*. Prove that*

$$p_2(\mathbf{x}) = \sum_{i=1}^{3} \left[ v_i L_i(2L_i - 1) + 4v_{i+3} L_{i+1} L_{i+2} \right], \ \forall \mathbf{x} \in K$$

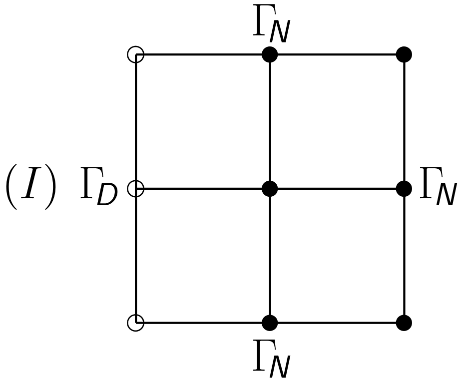
*where* $(L_1, L_2, L_3)$ *are the barycentric coordinates of* $\mathbf{x}$*, and*

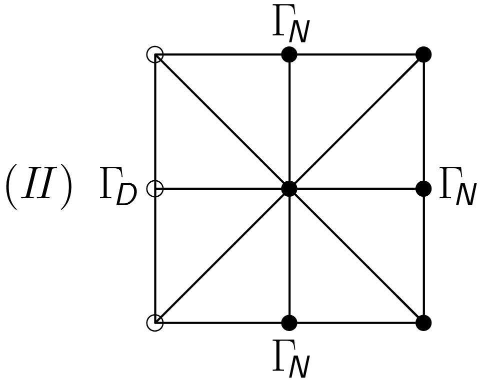
$$L_4 = L_1, L_5 = L_2, L_6 = L_3.$$

**Exercise 3.3** *Let* $\Omega = (-1, 1)^2$*. Construct for the problem:*

$$\begin{cases} \alpha u - \triangle u = 1, \ \forall \mathbf{x} \in \Omega \\ u|_{\Gamma_D} = 0 \\ \dfrac{\partial u}{\partial \mathbf{n}}|_{\Gamma_N} = 2 \end{cases}$$

*respectively a $\mathcal{Q}_1$-FEM based on the rectangular mesh $(I)$ and a $\mathcal{P}_1$-FEM based on the triangular mesh $(II)$:*




## Algorithmic properties

$$V_h = \{\varphi_1, \varphi_2, \cdots, \varphi_N\}.$$

$$AU = F,$$

with stiffness matrix $A = (a_{ij})_{i,j=1}^{N}, a_{ij} = a(\varphi_j, \varphi_i)$.

Properties of $A$:

- $A$ is positive definite if $a(\cdot, \cdot)$ is coercive
- $A$ is symmetric if $a(\cdot, \cdot)$ is symmetric
- $A$ is sparse, i.e., $a_{ij} = 0$ if supp$\varphi_i \cap$supp$\varphi_j = \varnothing$
- $cond(A) := cond_{sp}(A) := \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} = O(h^{-2})$.

# FD/FEM for Parabolic Equations

Consider time-dependent problem (IBVP):

$$\begin{cases} \frac{\partial u}{\partial t} - \triangle u = f & \forall \mathbf{x} \in \Omega \subset I\!R^2, \forall t \in (0, T) \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}) & \forall \mathbf{x} \in \Omega \\ u(\mathbf{x}, t)|_{\partial\Omega} = 0 & \forall t \in (0, T). \end{cases}$$

- Multiplying $v \in H_0^1(\Omega)$, and integrating

$$\int_\Omega \Big(\frac{\partial u}{\partial t} - \triangle u\Big) v d\mathbf{x} = (f, v).$$

- Integration by part

$$\Big(\frac{\partial u}{\partial t}, v\Big) + (\nabla u, \nabla v) = (f, v).$$

Variational formulation of (IBVP): find $u(\cdot, t) \in H_0^1(\Omega), t > 0$, such that

$$\begin{cases} \left(\frac{\partial u}{\partial t}, v\right) + (\nabla u, \nabla v) = (f, v), \ \forall v \in H_0^1(\Omega), \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}). \end{cases}$$

**Theorem 4.1 (ref to Theorem 11.1.1 in QV-p366)** *The energy estimate*

$$\|u(t)\|_0^2 + \int_0^t |u(\tau)|_1^2 \, d\tau \leqslant \|u_0\|_0^2 + \int_0^t \|f(\tau)\|_0^2 \, d\tau$$

*holds for each* $t \in [0, T]$.

**Theorem 4.2 (ref to Proposition 11.1.1 in QV-p370)**

$$\sup_{t \in (0,T)} |u(t)|_1^2 + \int_0^T \left\| \frac{\partial u}{\partial t}(t) \right\|_0^2 dt \leqslant 2(|u_0|_1^2 + \int_0^T \|f(t)\|_0^2 \, dt).$$

Galerkin method: find $u_h(\cdot, t) \in V_h \subset H_0^1(\Omega), t > 0$, such that

$$\begin{cases} \left(\frac{\partial u_h}{\partial t}, v_h\right) + (\nabla u_h, \nabla v_h) = (f, v_h), \ \forall v_h \in V_h, \\ u_h(\cdot, 0) = u_{0,h}, \end{cases}$$

where $u_{0,h}$ is an approximation of $u_0$ in $V_h$.

**Theorem 4.3 (ref to Proposition 11.2.1 in QV-p374)**

$$
\begin{aligned}
&\|u(t) - u_h(t)\|_0^2 + \int_0^t |u(\tau) - u_h(\tau)|_1^2 \, d\tau \\
&\leqslant \|u_0 - u_{0,h}\|_0^2 + Ch^2 \left( |u_{0,h}|_1^2 + |u_0|_1^2 + \int_0^t \|f(\tau)\|_0^2 \, d\tau \right).
\end{aligned}
$$

Proof. For each $t \in [0, T]$ define $e_h(t) := u(t) - u_h(t)$; from (11.1.5) and (11.2.1) one finds

$$
\frac{d}{dt}(e_h(t), v_h) + (\nabla e_h(t), \nabla v_h) = \left( \frac{\partial e_h}{\partial t}(t), v_h \right) + (\nabla e_h(t), \nabla v_h) = 0 \quad \forall\, v_h \in V_h
$$

For almost any fixed $t$, choose $v_h = u_h(t) - w_h, w_h \in V_h$ in (11.2.4). For each

$\varepsilon > 0$ and for almost any $t \in [0,T]$ we find

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}(e_h(t), e_h(t)) + (\nabla e_h(t), \nabla e_h(t)) \\
&= \left(\frac{\partial e_h}{\partial t}(t), u(t) - w_h\right) + (\nabla e_h(t), \nabla(u(t) - w_h)) \\
&\leqslant \left\|\frac{\partial e_h}{\partial t}(t)\right\|_0 \|u(t) - w_h\|_0 + |e_h(t)|_1 |u(t) - w_h|_1 \\
&\leqslant \left\|\frac{\partial e_h}{\partial t}(t)\right\|_0 \|u(t) - w_h\|_0 + \frac{1}{4\varepsilon}|u(t) - w_h|_1^2 + \varepsilon |e_h(t)|_1^2.
\end{aligned}$$

Choosing for almost any $t \in [0,T] w_h = \pi_h^k(u(t))$ (see (3.4.1)), we find

$$\|u(t) - \pi_h^k(u(t))\|_0^2 + h^2 |u(t) - \pi_h^k(u(t))|_1^2 \leqslant Ch^4 \|u(t)\|_2^2.$$

Integrating (11.2.5) in $(0,t)$ and choosing $\varepsilon = 1/2$ yields

$$\|e_h(t)\|_0^2 + \int_0^t |e_h(\tau)|_1^2$$
$$\leqslant \|u_0 - u_{0,h}\|_0^2 + Ch^2 \int_0^t \left( \left\| \frac{\partial u}{\partial t}(\tau) \right\|_0^2 + \left\| \frac{\partial u_h}{\partial t}(\tau) \right\|_0^2 + \|u(\tau)\|_2^2 \right).$$

As in (11.1.15) we have

$$\int_0^t \left\| \frac{\partial u_h}{\partial t}(\tau) \right\|_0^2 \leqslant C \left( |u_{0,h}|_1^2 + \int_0^t \|f(\tau)\|_0^2 \right).$$

The thesis follows from (11.1.19).

Let

$$V_h = \text{span}\{\varphi_1, \varphi_2, \cdots, \varphi_N\},$$

$$u_h(\mathbf{x},t)=\sum_{i=1}^{N}u_i(t)\varphi_i(\mathbf{x}),$$

then

$$\sum_{i=1}^{N}(\varphi_i,\varphi_j)\frac{du_i}{dt}+\sum_{i=1}^{N}a(\varphi_i,\varphi_j)u_i=(f,\varphi_j),\ \ j=1,2,\cdots,N.$$

Matrix statement

$$\begin{cases} M\dfrac{d\mathbf{u}}{dt}+A\mathbf{u}=\mathbf{f}, \\ \mathbf{u}(0)=\mathbf{u}_0, \end{cases}$$

where $m_{ji}=(\varphi_i,\varphi_j), a_{ji}=a(\varphi_i,\varphi_j)$.

Time discretization by a finite difference schema:

$$M\frac{\mathbf{u}^{m+1}-\mathbf{u}^m}{\triangle t}+A\mathbf{u}^m=\mathbf{f}^{m+1} \qquad \text{(Forward Euler)}$$

$$M\frac{\mathbf{u}^{m+1}-\mathbf{u}^m}{\triangle t}+A\mathbf{u}^{m+1}=\mathbf{f}^{m+1} \qquad \text{(Backward Euler)}$$

$$M\frac{\mathbf{u}^{m+1}-\mathbf{u}^m}{\triangle t}+A\frac{\mathbf{u}^{m+1}+\mathbf{u}^m}{\triangle t}=\mathbf{f}^{m+1/2} \qquad \text{(Cranck-Nicolson)}$$

Another way to discretize the parabolic equation

- First discretizing in time: $u^0 = u_0\ \forall \mathbf{x} \in \Omega$, compute $u^{m+1}$ for all $m = 0, 1, \cdots$ by

$$\begin{cases} \frac{u^{m+1}-u^m}{\triangle t} - \triangle u^{m+1} = f^{m+1},\ \forall \mathbf{x} \in \Omega \\ u^{m+1}|_{\partial\Omega} = 0,\ \forall \mathbf{x} \in \Omega. \end{cases}$$

(VP): find $u^{m+1} \in H_0^1(\Omega)$, such that

$$\frac{1}{\triangle t}(u^{m+1}, v) + (\nabla u^{m+1}, \nabla v) = \frac{1}{\triangle t}(u^m, v) + (f^{m+1}, v),\ \forall v \in H_0^1(\Omega).$$

-Then discretizing in space: find $u_h^{m+1} \in V_h \subset H_0^1(\Omega)$, such that

$$\frac{1}{\Delta t}(u_h^{m+1}, v_h) + (\nabla u_h^{m+1}, \nabla v_h) = \frac{1}{\Delta t}(u_h^m, v_h) + (f^{m+1}, v_h), \ \forall v_h \in V_h.$$

Stability with respect to the initial condition

- Backward Euler $\longrightarrow$ absolutely stable
- Forward Euler $\longrightarrow$ conditionally stable
- CN $\longrightarrow$ absolutely stable

**Exercise 4.1** *Analyze the stability of the CN schema:*

$$\frac{u^{m+1} - u^m}{\Delta t} - \frac{\triangle u^{m+1} + \triangle u^m}{2} = 0.$$

**Exercise 4.2** *Analyze the stability of the backward differentiation of second* *order (BD2):*

$$\frac{3u^{m+1} - 4u^m + u^{m-1}}{2\triangle t} - \triangle u^{m+1} = 0.$$

Error estimation (convergence): Backward Euler

Estimate the total error $u(\cdot, t^m) - u_h^m = u(\cdot, t^m) - u^m + u^m - u_h^m$

Let $e^m := u(\cdot, t^m) - u^m$, then

$$\frac{e^{m+1} - e^m}{\triangle t} - \triangle e^{m+1} = R^{m+1} = O(\triangle t).$$

Thus

$$\begin{aligned} \|e^{m+1}\|_0^2 &\leqslant (e^m, e^{m+1}) + \triangle t(R^{m+1}, e^{m+1}) \\ &\leqslant \|e^m + \triangle t R^{m+1}\|_0 \|e^{m+1}\|_0. \end{aligned}$$

$$
\begin{aligned}
\|e^{m+1}\|_0 &\leqslant \|e^m + \triangle t R^{m+1}\|_0 \\
&\leqslant \|e^m\|_0 + \triangle t\|R^{m+1}\|_0 \\
&\leqslant \|e^{m-1}\|_0 + \triangle t\|R^m\|_0 + \triangle t\|R^{m+1}\|_0 \\
&\leqslant \cdots \\
&\leqslant \|e^0\|_0 + \triangle t \textstyle\sum_{j=0}^{m+1} \triangle t\|R^j\|_0 \\
&\leqslant cM\triangle t^2 \\
&\leqslant cT\triangle t.
\end{aligned}
$$

Let $e_h^m := u^m - u_h^m$, then

$$\|e_h^m\|_0 \leqslant ch^{k+1}.$$

The total error

$$\|u(\cdot, t^m) - u_h^m\|_0 \leqslant c(T\triangle t + h^{k+1}).$$

**Exercise 4.3** *(Computing problem) Let* $\Omega = (0,1)^2$.

*1) Solve numerically by FD/*$\mathcal{P}_1$*-FEM method the problem:*

$$\begin{cases} \dfrac{\partial u}{\partial t} - \triangle u = f, \ (0,T)\times\Omega \\ u(\mathbf{x},0) = u_0(\mathbf{x}), \ \Omega \\ u(\mathbf{x},t)|_{\partial\Omega} = 0, \ (0,T) \end{cases}$$

*for* $f(\mathbf{x},t) = [8\pi^2\cos(t) - \sin(t)]\sin(2\pi x)\sin(2\pi y), u_0(\mathbf{x}) = \sin(2\pi x)\sin(2\pi y)$ *such that* $u(\mathbf{x},t) = \cos(t)\sin(2\pi x)\sin(2\pi y)$.

*2) Investigate the accuracy with respect to the time-step size* $\triangle t$ *and the mesh size* $h$.

# Summary

- FD/FEM methods for the parabolic equations
- FD methods for the parabolic equations

## Separated sections

- FD methods for ODEs: Euler schemes, trapezoidal, Leapfrog (midpoint), AB, AM, RK, and BDF etc.

    truncation error, stability, convergence

- FD methods for elliptic equations: centered schema (5-point), 9-point schema.

    truncation error, stability (energy estimates), convergence

- FEM methods for elliptic equations

- Distribution, derivative, Sobolev spaces ($L^2, H^1, H^m$ for example), norms, inner products, some inequalities, etc.

- Weak formulation

- Lax-Milgram lemma

- Galerkin method (error estimates)

- Finite element methods: mesh, space (piecewise polynomials space), basis functions, stiffness matrix, linear system, error estimates.

• FD/FEM methods for parabolic equations

**[Exercise]** Consider the transport-diffusion problem

$$\begin{aligned} u_t - u_{xx} + v u_x &= 0, \ \forall x \in (a,b), t \in (0,T) \\ u(a,t) = u(b,t) &= 0, \ t \in (0,T) \\ u(x,0) &= u_0(x), \ \forall x \in (a,b) \end{aligned}$$

where $v$ is a constant. Analyze the two following methods

1. Finite difference schema, if $v \geqslant 0$,

$$\begin{aligned}\frac{u_i^{n+1}-u_i^n}{k}-\frac{u_{i+1}^{n+1}-2u_i^{n+1}+u_{i-1}^{n+1}}{h^2}+v\frac{u_i^{n+1}-u_{i-1}^{n+1}}{h} &= 0,\ \forall i=1,\cdots,N-1,\\ u_0^{n+1}=u_N^{n+1} &= 0,\\ u^0 &= u_0,\end{aligned}$$

or, if $v \leqslant 0$,

$$\begin{aligned}\frac{u_i^{n+1}-u_i^n}{k}-\frac{u_{i+1}^{n+1}-2u_i^{n+1}+u_{i-1}^{n+1}}{h^2}+v\frac{u_{i+1}^{n+1}-u_i^{n+1}}{h} &= 0,\ \forall i=1,\cdots,N-1,\\ u_0^{n+1}=u_N^{n+1} &= 0,\\ u^0 &= u_0,\end{aligned}$$

in an uniform mesh $\{x_i\}_{i=0}^N, x_i=a+ih, h=(b-a)/N$, $\{t^n\}_{n=0}^M, t^n=nk, k=T/M$.

2. Finite element method: find $u_h^{n+1} \in V_h$, such that

$$\begin{aligned} (\frac{u_h^{n+1} - u_h^n}{k}, v_h) + (u_{h,x}^{n+1}, v_{h,x}) + v(u_{h,x}^{n+1}, v_h) &= 0, \ \forall v_h \in V_h, \\ u_h^0 &= u_0, \end{aligned}$$

where $V_h$ is the space of piecewise linear functions based on the mesh $\{x_i\}_{i=0}^N, x_i = a + ih, h = (b-a)/N$.

3. How about 2D case?